



Rendering by Manifold Hopping

HEUNG-YEUNG SHUM, LIFENG WANG, JIN-XIANG CHAI AND XIN TONG

Microsoft Research, Asia, 49 Zhichun Road, Beijing 100080, P.R. China

hshum@microsoft.com

Received July 3, 2001; Revised February 13, 2002; Accepted March 13, 2002

Abstract. In this paper, we present a novel image-based rendering technique, which we call *manifold hopping*. Our technique provides users with *perceptually continuous navigation* by using only a small number of strategically sampled manifold mosaics or multiperspective panoramas. Manifold hopping has two modes of navigation: moving continuously along any manifold, and discretely between manifolds. An important feature of manifold hopping is that significant data reduction can be achieved without sacrificing output visual fidelity, by carefully adjusting the hopping intervals. A novel view along the manifold is rendered by locally warping a single manifold mosaic using a constant depth assumption, without the need for accurate depth or feature correspondence. The rendering errors caused by manifold hopping can be analyzed in the signed Hough ray space. Experiments with real data demonstrate that we can navigate smoothly in a virtual environment with as little as 88k data compressed from 11 concentric mosaics.

Keywords: manifold mosaic, concentric mosaic, perceptually smooth navigation, warping, plenoptic functions, image-based rendering, video computing

1. Introduction

Image-based rendering (IBR) in general simulates a continuous range of camera viewpoints from a discrete set of input images (McMillan and Bishop, 1995; Kang, 1999). Much of the work on image-based rendering has depended on view interpolation (e.g., Chen and Williams, 1993; McMillan and Bishop, 1995; Seitz and Dyer, 1996; Avidan and Shashua, 1997). View interpolation is, however, a difficult task because feature correspondence needs to be accurately established. Correspondence is particularly difficult if it involves two images which are taken from distant viewpoints.

When accurate depth information of an image is available, the image can be rendered from a nearby view by directly warping the input image according to its depth (Mark et al., 1997; McMillan, 1999). To deal with the occlusion problems, Layered depth images (LDI) (Shade et al., 1998) maintain multiple depth values for each pixel in a single image so that warping

may be computed with a single image. The sampling issues in LDI are considered in an LDI tree (Chang et al., 1999) by adaptively selecting an LDI for each pixel. However, recovering precise depth information is still very difficult.

Many IBR techniques avoid the difficult correspondence problem or the need for accurate depth information by employing a substantial number of images. Techniques such as Light Field (Levoy and Hanrahan, 1996), Lumigraph (Gortler et al., 1996), and Concentric Mosaics (Shum and He, 1999) densely sample rays in the space based on the plenoptic function (Adelson and Bergen, 1991). To render an image at any novel viewpoint using these techniques, nearby rays are chosen and interpolated from the sampled plenoptic functions. For example, concentric mosaics are constructed from a dense sequence of video images, and constant depth assumptions need to be made about the scene depth in order to locate the best “nearby” rays for optimal rendering quality. Applications of light

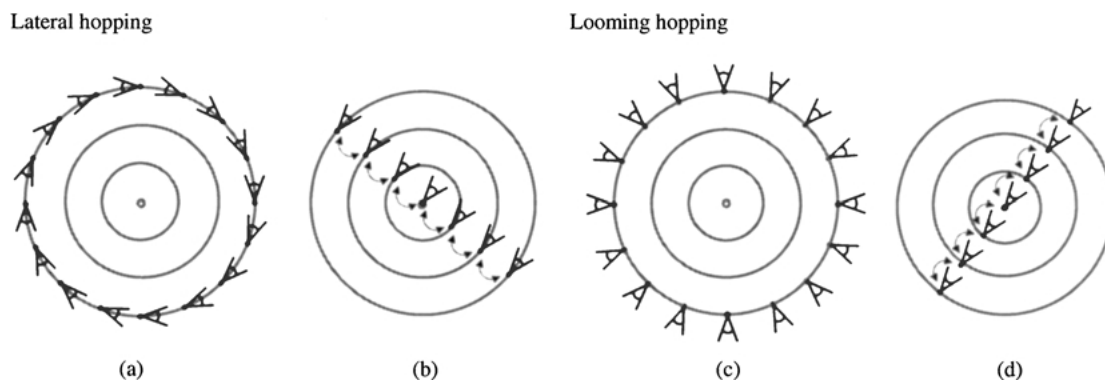


Figure 1. Manifold hopping using concentric mosaics: A plan view. Manifold hopping has two modes of navigation: (a) and (c) move continuously along any manifold, and (b) and (d) discretely across manifolds. The arrows in (b) and (d) indicate that the user can only hop to the viewpoints on the circle, but not stop anywhere in between. Two classes of manifold hopping are shown here: Lateral hopping whose discrete mode of navigation is perpendicular to the viewing direction, and looming hopping whose discrete mode of navigation is along the viewing direction. Lateral hopping uses tangent concentric mosaics (Fig. 2(a)), while looming hopping employs normal concentric mosaics (Fig. 2(b)). Note that the rendering views are restricted on the circles. Rendering images have limited field of views.

fields are limited because of the large amount of data needed.

An effective way to reduce the amount of data needed for IBR is to constrain the motion or the viewpoints of the rendering camera. For example, the movie-map system (Lippman, 1980) and the QuickTime VR system (Chen, 1995) allow a user to explore a large environment only at pre-specified locations. Even though a continuous change in viewing directions at each node is allowed, these systems can only jump between two nodes that are far apart, thus causing visual discontinuity and discomfort to the user. However, perceived continuous camera movement is very important for a user to smoothly navigate in a virtual environment. Recently, several panoramic video systems have been built to provide a dynamic and immersive “video” experience by employing a large number of panoramic images.

In this paper, we propose a novel image-based rendering technique, which we call *manifold hopping*. In our work, scene appearance is represented by a collection of manifold mosaics. A manifold mosaic (Peleg and Herman, 1997), similar to a multiperspective panorama (Wood et al., 1997), is assembled from rays that are captured from multiple viewpoints (Rademacher, 1998; Wood et al., 1997; Gupta and Hartley, 1997; Zheng and Tsuji, 1990). Furthermore, we assume in this paper that a manifold mosaic has a one-to-one mapping between each pixel and its corresponding scene point. The term manifold hopping is used to indicate that while motion is continuous within

a manifold, motion between manifolds is discrete, as shown in Fig. 1.

Manifold hopping significantly reduces the amount of input data without sacrificing output visual quality, by employing only a small number of strategically sampled manifold mosaics. Our technique is based on the observation that, for human visual systems to perceive continuous motion, it is not essential to render novel views at infinitesimal steps. Moreover, manifold hopping does not require accurate depth information or correspondence between images. At any point on a given manifold, a novel view is generated by locally warping the manifold mosaic with a constant depth assumption, rather than interpolating from two or more mosaics. Although warping errors are inevitable because the true geometry is unknown, local warping does not introduce structural features such as double images which can be visually disturbing.

The remainder of this paper is organized as follows. After reviewing the concepts of manifold mosaics and concentric mosaics, we present an overview of manifold hopping in Section 2. Detailed analysis of manifold hopping (with lateral movements) is presented in Section 3. In particular, using the signed Hough ray space, the hopping interval and field of view are analyzed for radial hopping using concentric mosaics. In Section 4, manifold hopping with looming movements is discussed in terms of the extended signed Hough ray space. In Section 5, we use parallel mosaics for hopping from outside. Continuous close-up views of objects can be seen using parallel mosaics, while hopping around

the objects. Experiments using real and synthetic images are shown in Section 6. Other hopping choices are discussed in Section 7. Finally, we conclude and propose future research directions in Section 8.

2. Overview of Manifold Hopping

In this section, we introduce manifold mosaics, view interpolation using manifold mosaics, warping manifold mosaics, and manifold hopping. Throughout this section, concentric mosaics are used as examples of manifold mosaics to illustrate these concepts.

2.1. Manifold Mosaics

A multiperspective image is assembled from rays are captured from multiple viewpoints (e.g., Zheng and Tsuji, 1990). Multiperspective images have also been called MCOP images (Rademacher, 1998), multiperspective panoramas (Wood et al., 1997), pushbroom images (Gupta and Hartley, 1997), and manifold mosaics (Peleg and Herman, 1997), among other names. In this paper, we define a *manifold mosaic* as a multiperspective image where each pixel has a one-to-one mapping with a scene point.¹ Therefore, a conventional perspective image, or a single perspective panorama, can be regarded as a degenerate manifold mosaic in which all rays are captured at the same viewpoint.

We adopt the term manifold mosaic from Peleg and Herman (1997) because the viewpoints are generally taken along a continuous path or a manifold (surface or curve). For example, concentric mosaics are manifold mosaics constructed from rays taken along concentric circles (Shum and He, 1999). Specifically, at each point on a circle, a slit image of single pixel width is taken. By assembling all slit images captured along a circle, a concentric mosaic is formed. Two kinds of concentric mosaics are shown in Fig. 2 where rays are taken in the *tangential* direction (Fig. 2(a)), and in the *normal* direction (Fig. 2(b)), respectively. Concentric mosaics constitute a 3D plenoptic function because they are sampled naturally by three parameters: rotation angle, radius, and vertical elevation. Clearly there is a one-to-one mapping between pixels in a concentric mosaic and their corresponding scene points.

Although many previous image-based rendering techniques (e.g., view interpolation, 3D warping, etc.) are developed for perspective images, they can be applied to manifold mosaics as well. For example, 3D

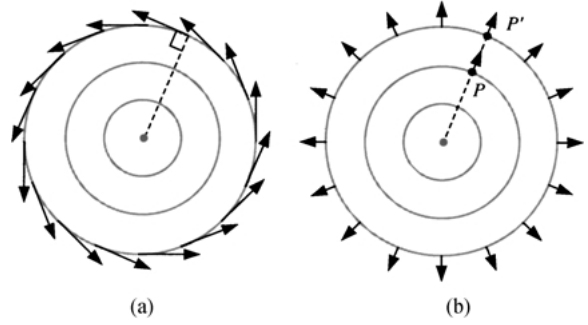


Figure 2. Capturing two kinds of concentric mosaics: A plan view. A concentric mosaic is assembled by unit width slit images (a) tangent to the circle; and (b) normal to the circle. We call (a) tangent concentric mosaics and (b) normal concentric mosaics. Tangent concentric mosaics are called concentric mosaics in Shum and He (1999).

warping has been used to reproject a multiple-center-of-projection (MCOP) image in Rademacher (1998) and Oliveira and Bishop (1999) where each pixel of an MCOP image has an associated depth. Stereo reconstruction from multiperspective panoramas has also been studied (Shum and Szeliski, 1999).

It has been shown (Shum and He, 1999) that a novel view inside the capturing circle can be rendered from the concentric mosaics without any knowledge about the depth of the scene. From densely sampled concentric mosaics, a novel view image can be rendered by linearly interpolating nearby rays from two neighboring concentric mosaics. In addition, a constant depth is assumed to find the best “nearby” rays for optimal rendering quality (Chai et al., 2000b). Figure 3(a) illustrates a

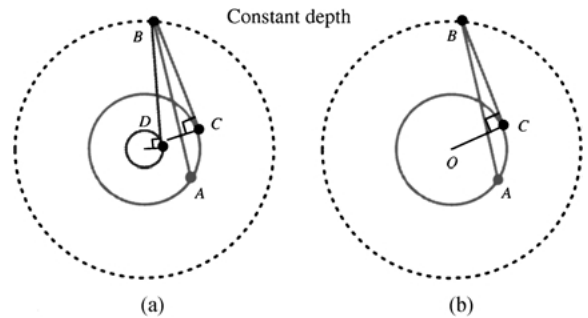


Figure 3. Rendering concentric mosaics with (a) view interpolation and (b) local warping. (a) A ray from viewpoint A is projected to the constant depth surface (represented as a dotted circle) at B, and interpolated by two rays BC and BD that are retrieved from neighboring concentric mosaics. (b) A ray from viewpoint A is projected to the constant depth surface at B, and reprojected to the nearest concentric mosaic by the ray BC.

rendering ray that is interpolated by two rays captured in nearby concentric mosaics. Despite the inevitable vertical distortion, concentric mosaics are very useful for wandering around (on a plane) in a virtual environment. In particular, concentric mosaics are easy to capture by using an off-centered video camera rotating along a circle.

2.2. Warping Manifold Mosaics

View interpolation can create high quality rendering results when the sampling rate is higher than Nyquist frequency for plenoptic function reconstruction (Chai et al., 2000b). However, if the sampling interval between successive camera locations is too large, view interpolation will cause aliasing artifacts, creating double images in the rendered image. Such artifacts can be reduced by the use of geometric information (e.g., Gortler et al., 1996; Chai et al., 2000b), or by pre-filtering the light fields (Levoy and Hanrahan, 1996; Chai et al., 2000b) (thus reducing output resolution).

A different approach is to locally warp manifold mosaics, which is similar to 3D warping of a perspective image. An example of locally warping concentric mosaics using an assumed constant depth is illustrated in Fig. 3(b). Any rendering ray that is not directly available from a concentric mosaic (i.e., not tangent to a concentric circle) can be retrieved by first projecting it to the constant depth surface, and then re-projecting it

to the concentric mosaic. Therefore, a novel view image can be warped using the local rays captured on a single concentric mosaic, rather than interpolated by collecting rays from two or more concentric mosaics.

For humans to perceive a picture correctly, it is essential that the image of an object should not contain any structural features that are not present in the object itself (Zorin and Barr, 1995). Double images, which are common artifacts from view interpolation with poor correspondence, unfortunately result in mistakenly perceived structural features in the observed objects, e.g., more noticeable edges. On the other hand, locally warping a multiperspective image preserves structural features. An example of locally warping a concentric mosaic is shown in Fig. 4, with images of different FOV's. The projection error in the rendered image caused by warping the concentric mosaic with (incorrect) constant depth assumption increases as the field of view becomes larger. Note the distortion toward the right edge in Fig. 4(b). The geometric distortions introduced by local warping methods because of imprecise geometric information are, however, tolerated by human visual perception when the field of view (FOV) of the rendering image is small (e.g., Fig. 4(a)).

2.3. Hopping Classification and Issues

We now introduce the idea of manifold hopping using a small number of concentric mosaics to observe

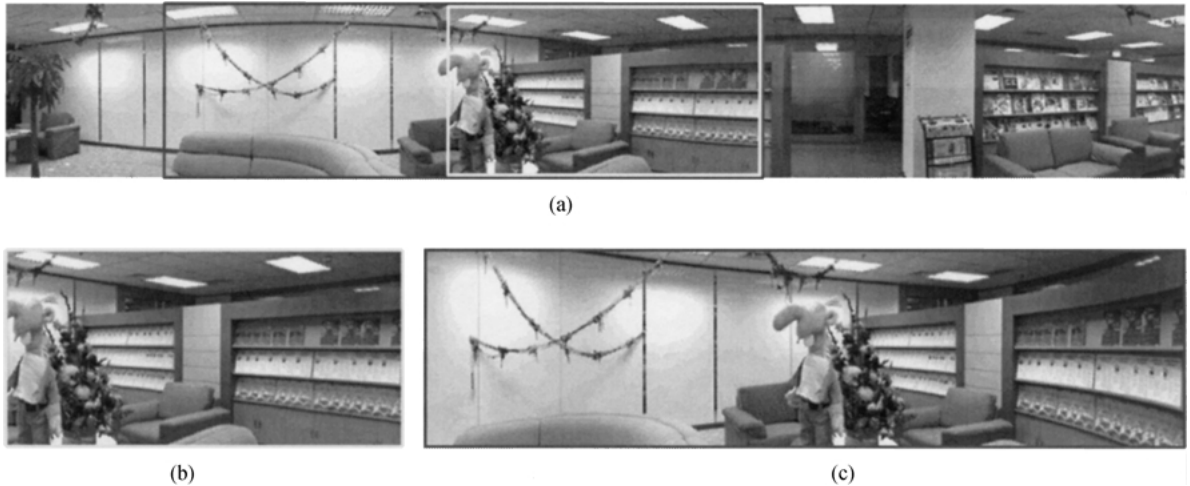


Figure 4. Local warping with an assumed constant depth: (a) part of a concentric mosaic; (b) a rendered view with FOV = 45; and (c) another rendered view with FOV = 90. The distortion error towards the right edge of (c) can be clearly seen as straight lines become curved. The image is rendered column by column with local warping. Note that in (c), the vertical field of view is significantly reduced as some of the ceiling lights become invisible.

an environment from the inside looking out. Manifold hopping has two modes of navigation: moving continuously along any of the concentric circles as shown in Fig. 1(a) and (c), but discretely along the radial direction as in Fig. 1(b) and (d).

Manifold hopping works because moving continuously along any concentric circle uses local warping, which preserves structural features. In addition, moving discretely along the radial direction can be made perceptually smooth if the interval can be made reasonably small. A key observation is that there exists a critical hopping interval for users to perceive a smooth navigation. In other words, manifold hopping is able to provide users with *perceived continuous camera movement*, without continuously rendering viewpoints at infinitesimal steps. As a result, manifold hopping significantly reduces the input data size without accurate depth information or correspondence.

Figure 1(a) also shows that, at any point on a circle, the rendering view is constrained to be on the circle and the viewing direction along the tangent line to minimize the rendering errors caused by local warping. Note that no parallax is observed from these views generated on the same circle using the same concentric mosaic. Parallax and lighting changes are captured in manifold hopping because of the viewpoint variations across different concentric circles, as shown in Fig. 1(b).

In this paper, we describe two types of manifold hopping with concentric mosaics: lateral hopping, whose discrete mode of navigation (Fig. 1(b)) is perpendicular to the viewing direction; and looming hopping, whose discrete mode of navigation (Fig. 1(d)) is along the viewing direction. Note that for each type of hopping, there are two modes of navigation, namely the continuous mode along the manifold and discrete mode between manifolds. The type of hopping is named after the direction of discrete navigation.

Detailed analysis of manifold hopping is needed to address the following important questions.

- What is the largest field of view that still produces acceptable local warping error?
- How large can the hopping interval be so that continuous motion can be perceived?

These questions are answered in detail for lateral hopping in the next section after we introduce the signed Hough ray space. In Section 4, we introduce the extended signed Hough ray space to analyze looming hopping.

3. Analysis of Lateral Hopping Using the Signed Hough Ray Space

The Hough transform is known to be a good representation for lines. However, it is not suitable for representing rays that are directional. The conventional Hough space can be augmented to a signed Hough ray space (Chai et al., 2000a), or an oriented line representation (Levoy and Hanrahan, 1996), by using the following right-hand rule: a ray that is directed in a counter-clockwise fashion about the coordinate center is labeled positive, otherwise is labeled negative. A “positive” ray is represented by (r, θ) , whereas its “negative” counterpart is $(-r, \theta)$ where r is always a positive number. Figure 5 shows four different rays in a 2D space and their corresponding points in the signed Hough space.

Figure 6 shows three typical viewing setups and their representations in the signed Hough space. For example, a panoramic image (i.e., rays collected at a fixed viewpoint in Cartesian space) is represented as a sampled sinusoidal curve in the parameter space, located at (r_0, θ_0) as shown in Fig. 6(a). A concentric mosaic shown in Fig. 6(b) is mapped to a horizontal line, whereas parallel projection rays (Fig. 6(c)) are mapped to a vertical line in the signed Hough space. Thus, captured perspective images can be easily transformed into samples in the parameter space. Rendering a novel view in the scene is equivalent to extracting a partial or complete sinusoidal curve from the signed Hough space.

When the hopping direction is perpendicular to the viewing direction, as shown in Fig. 1, we call it lateral hopping. In the signed Hough space, such a hopping is illustrated in Fig. 7 where a segment of a sinusoidal curve is approximated by a line segment. Equivalently, at each rendering viewpoint, a perspective

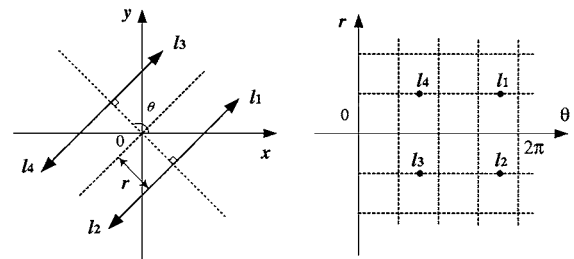


Figure 5. Definition of the signed Hough ray space: Each oriented ray in Cartesian space at the left is represented by a sampled point in the signed Hough space on the right.

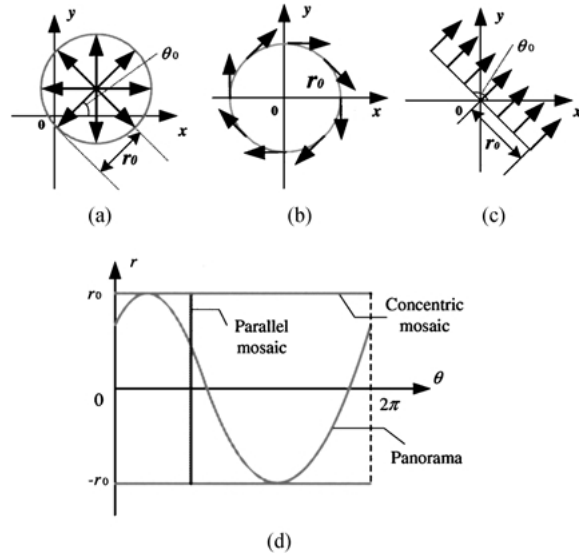


Figure 6. Three typical viewing setups and their respective sampled curves in the signed Hough space: (a) a panorama at a fixed point; (b) a concentric mosaic; (c) a parallel projection mosaic; and (d) their respective sampled curves in the signed Hough space. Two concentric mosaics (straight lines at r_0 and $-r_0$) are shown in (d) to represent rays captured at opposite directions along the circle. Note that a perspective image is only part of a panorama, thus represented by a segment of a sinusoidal curve in the signed Hough space.

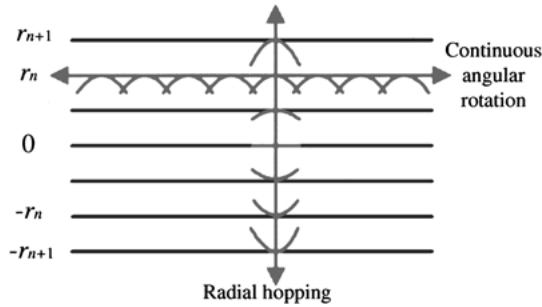


Figure 7. Hopping between concentric mosaics along a radial direction in the signed Hough space. Continuous rotation is achieved along any of the concentric circles, but hopping is necessary across any radial direction. In radial hopping, the curve segment varies from r_n to r_{n+1} because the corresponding sine curve is different, as shown in Fig. 6.

image is approximated by part of a concentric mosaic.

Obviously, the smaller the hopping interval, the smaller the rendering error. On the other hand, the larger the hopping interval, the less data needed for wandering around an environment. We argue that a fairly large hopping interval for manifold hopping can be perceptually acceptable.

3.1. When is Local Warping Good Enough?

When moving on a concentric mosaic, the horizontal field of view should be constrained within a certain range so that the distortion error introduced in local warping from a multiperspective image to a perspective image will not cause much visual discomfort to the user.

The distortion threshold η_d is defined as the maximum allowable distance between point A and point B in Fig. 8. These two points are projections of the rightmost pixel that are locally warped with distance R_1 (assumed distance) and R_2 (corrected distance), respectively. A radial hopping camera must satisfy the following:

$$\Delta\theta = \theta_B - \theta_A \leq \eta_d, \quad (1)$$

where $\theta_B = \sin^{-1} \frac{r_n}{R_2} - \sin^{-1} \frac{r_n - \Delta r}{R_2}$ is the angular difference when the object at R_2 distance viewed along circles of r_n and $r_n - \Delta r$. θ_A is defined similarly for object located at R_1 . Thus,

$$\begin{aligned} & \sin^{-1} \frac{r_n}{R_2} - \sin^{-1} \frac{r_n - \Delta r}{R_2} - \sin^{-1} \frac{r_n}{R_1} \\ & + \sin^{-1} \frac{r_n - \Delta r}{R_1} \leq \eta_d \end{aligned} \quad (2)$$

If parallel interpolation is applied to local warping by assuming the depth R_1 at infinity, we can simplify the above constraint to

$$\sin^{-1} \frac{r_n}{R_2} - \sin^{-1} \frac{r_n - \Delta r}{R_2} \leq \eta_d \quad (3)$$

$$1 - \cos\left(\frac{Fov}{2}\right) = \frac{\Delta r}{r_n} \quad (4)$$

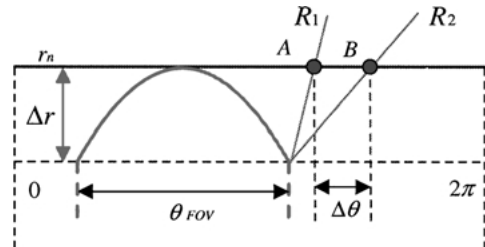


Figure 8. Analysis of maximum FOV: Warping error due to the incorrect depth value.

From the above two equations, we can derive the maximum FOV under parallel interpolation as

$$\cos\left(\frac{FOV}{2}\right) \geq \cos \eta_d - \sqrt{\left(\frac{R_2}{r_n}\right)^2 - 1} \sin \eta_d \quad (5)$$

The above equation shows that, under parallel interpolation, the maximum FOV for a hopping camera depends on the radius of the concentric mosaic, the scene depth, and the distortion error threshold. The field of view can be significantly increased when the object moves farther away. A smaller radius enables a larger FOV. For example, a panorama with a very large FOV can be rendered as the radius goes to zero. In addition, warping with constant depth (rather than infinite depth) can further increase the maximum FOV.

Consider a scene that is located along a circle whose radius is four times that of the outermost concentric mosaic. If we assume that the distortion threshold is 1° (that is a flow of 5 pixels for a mosaic with width 1800), the maximum allowable FOV is 42.42° .

Fortunately human visual perception does not require a very large field of view for a hopping camera when wandering in a virtual environment. It has also been shown that 36° is close to perceptually optimal for most people (Zorin and Barr, 1995). It is well known that small FOV perspective images are generated from a large multiperspective panorama for the purpose of animation (Wood et al., 1997).

3.2. How Large Can the Hopping Interval Be?

The efficiency and effectiveness of hopping depend on the size of sample intervals along both the radial and angular directions. The angular direction is sampled uniformly and densely to ensure a continuous rotation. The maximum hopping interval Δr allowed for smooth visual perception is determined by the threshold of the horizontal pixel flow D_0 (in angular measurement) between two neighboring frames. The analysis of vertical parallax is ignored in our analysis due to the nearly horizontal epipolar geometry between neighboring concentric mosaics (Shum and Szeliski, 1999).

Suppose that a point at a distance R_0 is seen in two concentric mosaics r_n and r_{n+1} , respectively. As shown in Fig. 9, the horizontal parallax $\Delta\theta$ between two observed pixels A and B at the two concentric mosaics

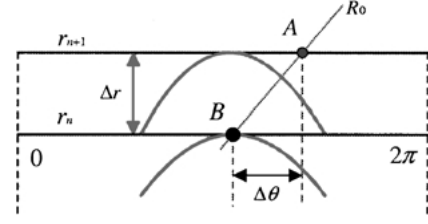


Figure 9. Analysis of hopping size: Horizontal parallax change due to viewpoint change.

satisfies

$$\Delta\theta = \sin^{-1}\left(\frac{r_n + \Delta r}{R_0}\right) - \sin^{-1}\left(\frac{r_n}{R_0}\right) \leq D_0 \quad (6)$$

which leads to the maximum hopping size

$$\Delta r = \sqrt{R_0^2 - r_n^2} \sin D_0 + r_n \cos D_0 - r_n \quad (7)$$

$$= R_0 \sin\left(D_0 + \sin^{-1}\frac{r_n}{R_0}\right) - r_n. \quad (8)$$

The above equation reveals that the sample interval along the radial direction depends on the depth (R_0), the smooth perception threshold (D_0), and the radius (r_n) of the concentric mosaic. Specifically, we observe:

- Sampling along the radial direction is nonlinear. The smaller the radius, the larger the hopping intervals should be.
- The hopping interval can be increased with object distance. When objects are located at infinity, all concentric mosaics degenerate to the same panorama.
- A larger threshold D_0 allows for a larger hopping interval along the radial direction. As $\Delta r \rightarrow 0$, the hopping interval $D_0 \rightarrow 0$. This is equivalent to rendering with concentric mosaics (Shum and He, 1999). On the other hand, if it is not required to observe parallax, a single manifold mosaic is enough for a user to look at any viewing direction.

The choice of threshold D_0 is closely related to the human visual system. It is well known that, for a human to observe smoothly moving pictures, the frame rate is 24 frames/second. Suppose that the average speed of rotation for a person to observe an environment is below 48° /second, then D_0 should be 2° . In other words, a person can tolerate 2° of average pixel flow for two neighboring frames and still observe smooth and continuous motion.

Consider a particular scene in which the radius of the outermost concentric mosaic is 1 unit and the objects are located at a distance of 4 units. If D_0 is 1.5° , we have $\Delta r = 0.1$. Therefore, we need only 21 concentric mosaics (two for each concentric circle and one for the center). This is a significant reduction from 320 rebinned concentric mosaics needed in rendering with concentric mosaics (Shum and He, 1999).

4. Analysis of Looming Hopping Using the Extended Signed Hough Space

In the previous section, we have analyzed manifold hopping where the hopping direction is perpendicular to the viewing direction. If the hopping direction is along the viewing direction, i.e., if the user moves forward and backward, we cannot use the conventional concentric mosaics assembled by rays along the tangent lines of the concentric circle. Instead, hopping with a looming motion can be achieved if we construct *normal* concentric mosaics that are formed by slit images with unit pixel width along the normal directions of the concentric circle, as shown in Fig. 2(b). A novel view at any point on a circle can be rendered by locally warping rays from the normal concentric mosaic near the viewpoint, as shown in Fig. 1(c).

The signed ray space is no longer adequate for analyzing looming hopping. For a looming motion, we need to represent points along the same ray differently. Therefore, we introduce the extended signed Hough space, defined by a 3-tuple (r, θ, d) where d is the distance from the origin to the location where the ray is captured. Two points (P and P') along the same ray have identical (r, θ) but different values of d , as shown in Fig. 10. And d will take the same sign as r to differ-

entiate a “positive” ray from a “negative” one, similar to the signed Hough space. Although rays captured at P and P' are the same in the plan view of Fig. 2(b), slit images captured at these two points are different.

Figure 10 also shows three different mosaics represented in the extended Hough space.

- A panorama: $r = d \sin(\theta - \phi)$;
- A tangent concentric mosaic: $r = d = r_n$;
- A normal concentric mosaic: $r = 0$ and $d = r_n$.

Note that ϕ is the constant angle for the viewpoint, and r_n is the diameter of one of the concentric circles. It becomes evident now why the location of the ray, which was ignored in lateral hopping (in the signed Hough space), should be considered in looming hopping because r is always zero under looming. Therefore, the (r, θ, d) representation is necessary and sufficient to index rays in 2D (plan view in Fig. 2(b)) to capture the looming effect as the user moves forward and backward.

Figure 11 illustrates looming hopping in the extended signed Hough space. Similar to lateral hopping in the signed Hough space (Fig. 7), rendering a novel view in looming hopping is also equivalent to approximating a partial sinusoidal curve by a line segment of a normal concentric mosaic. Unlike lateral hopping, however, each sinusoidal curve is constructed at a different d . For clarity, we skip the looming hopping interval analysis in the extended signed Hough space, which is similar to the analysis in the signed Hough space in the previous section.

Lateral hopping is also illustrated in Fig. 11. In the (r, θ, d) space, the plane for lateral hopping is $r = d$, but $r = 0$ for looming hopping. The sinusoidal curve

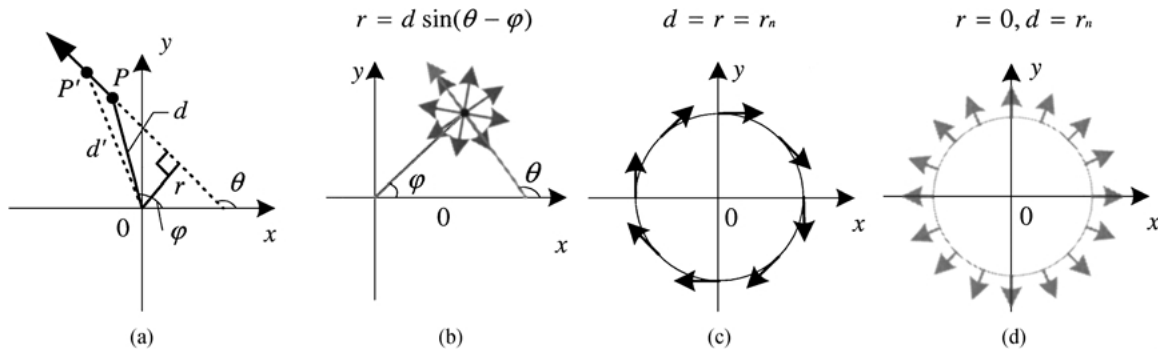


Figure 10. (a) The extended signed Hough ray space is defined by three parameters (r, θ, d) . Different points on the same ray have different d values. (b)–(d) A panorama, a tangent concentric mosaic, and a normal concentric mosaic are represented in the extended signed Hough space.

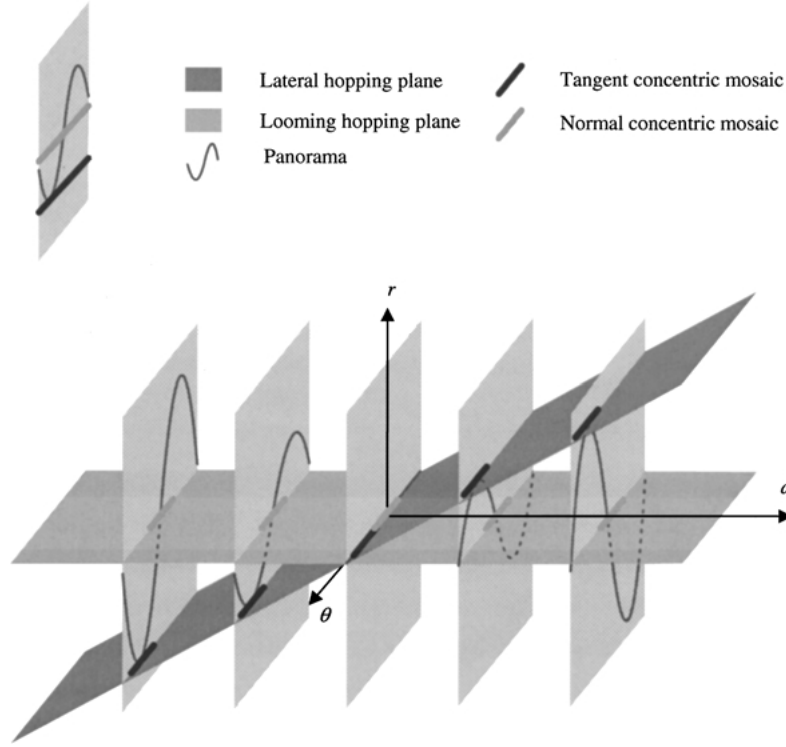


Figure 11. Looming hopping with normal concentric mosaics, and lateral hopping with tangent concentric mosaics in the extended signed Hough space. Rendering a novel perspective view is equivalent to approximating a sinusoidal curve segment by a straight line segment representing part of a concentric mosaic. In looming hopping, green segments are used to approximate the sine curve at different d values on the brown $r = 0$ plane. In lateral hopping, black segments are used to approximate the sine curve at different r (and d) values on the blue $r = d$ plane.

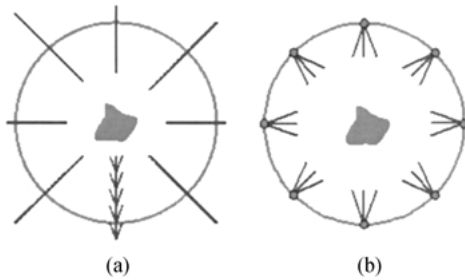


Figure 12. Hopping from outside: (a) translating continuously in the radial direction (toward the object); and (b) hopping discretely in the angular direction (around the object).

segment is approximated around the maximum r in lateral hopping, and around $r = 0$ for looming hopping. If we project the lateral hopping plane in (r, θ, d) space onto the $d = 0$ plane, we obtain the (r, θ) counterpart for lateral hopping. There is therefore a duality between lateral hopping (r, θ) and looming hopping (d, θ) .

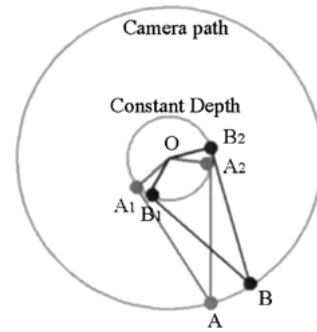


Figure 13. Hopping interval between two perspective images (viewed from outside the object) is related to the field of view of the object (A_1A_2). The object is assumed to be bounded within the constant depth circle.

5. Hopping from Outside

Concentric mosaics are suitable for wandering around in an environment when a user is looking outwards. When the user is looking at an object, it is desirable

to observe the object from outside at different viewing angles. In addition, it is important to have close-up views.

For simplicity of analysis and capturing, we consider camera motion on a plane as an object rotates in front of a camera. A sequence of perspective images are then taken along the camera path (i.e., a circle). We also constrain the rendering camera to have continuous motion along the radial direction (moving towards and away from the object along a line) as shown in Fig. 12(a), and discrete hopping motion in the angular direction as shown in Fig. 12(b).

5.1. Hopping Between Perspective Images

This rendering camera motion can be achieved by simply using the perspective images captured in the original sequence. Assuming a constant depth for the object, we can reproject perspective images to any novel views along the radial direction. However, only zooming effect, not the parallax effect, can be observed when the camera moves along the radial direction. When the camera moves away from the object, we can not observe any additional part of the object around the boundary other than what is in the original image.

5.1.1. Angular Hopping Interval. Many previous systems have used multiple images to observe a single object from outside. It is, however, important to study how large the hopping interval should be to ensure a perceived smooth transition between the images.

As shown in Fig. 13, two neighboring cameras A and B are located along the circle (with radius R) of camera path. The object is assumed to be at the circle (with radius r) of constant depth. $OA = OB = R$, and $OA_1 = OB_1 = r$. The camera spacing is $\alpha = AOB = A_1OB_1$. Let $\beta = A_1AO$, $2\beta = A_1AA_2$, and $\sin \beta = r/R$. The angular flow between two images can be approximated as

$$\Delta\theta \approx 2\beta \frac{\widehat{AB}}{A_1A_2} = 2\beta \frac{\alpha}{\pi - 2\beta}. \quad (9)$$

Therefore, given the pixel flow threshold D_0 , we obtain the camera spacing as

$$\alpha = \left(\frac{\pi}{2\beta} - 1 \right) D_0. \quad (10)$$

For example, if D_0 is 1° , and $R = 3r$, then α is computed as 4° . In other words, we need to capture 90 images along the circle.

5.2. Hopping Between Parallel Projection Mosaics

Another way to achieve continuous radial motion is to use parallel projection mosaics. Parallel mosaics are formed by collecting all parallel rays in the same direction. We call this angular hopping with parallel mosaics.

Because parallel projection cameras are not commonly available, we rebin parallel mosaics by taking the parallel rays from a dense sequence of perspective images taken along a circle outside the object. Figure 15 shows a projective image from the original sequence and the rebinned parallel mosaic. Note that the rebinned mosaic is called 1D parallel mosaic because the vertical direction is still perspective, only the horizontal direction is under parallel projection.

Assuming a constant depth for the object, we can reproject parallel mosaics to any novel view along the radial direction, as shown in Fig. 14. Warping 1D parallel mosaic in Fig. 15(b) using constant depth is shown in Fig. 15(c). Even though warping errors are created, such as those around the boundary of the object, they are small enough to cause little visual distortion.

5.2.1. Close-Up Views. Rendering a novel view with angular hopping using parallel mosaics can again be explained in the signed Hough space. Continuous motion along the angular direction is obtained by

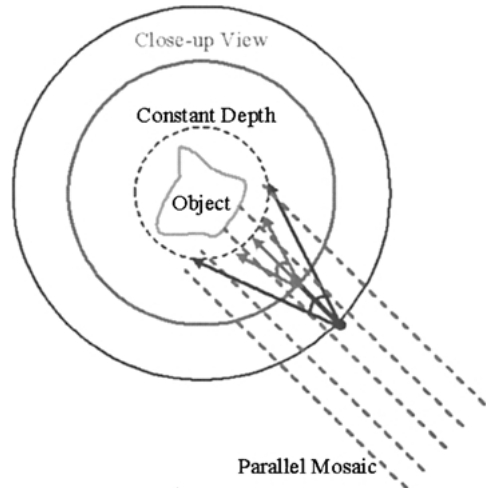


Figure 14. Reprojecting a parallel mosaic (shown as parallel dotted green lines) to different perspective images along the radial direction using constant depth assumption. The image shown by blue arrows is viewed at a normal distance away, while the image with red arrows is a close-up view.

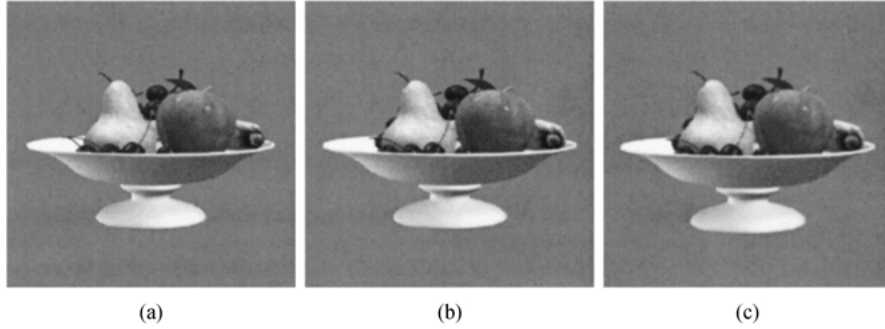


Figure 15. Warping parallel projection images: (a) a perspective image; (b) a 1D parallel projection mosaic; and (c) 1D mosaic of (b) warped with constant depth. (c) and (a) are mostly similar except around edges (e.g., on the left edge of the pear). An advantage of using parallel mosaics is to have higher resolution especially for close-up views.

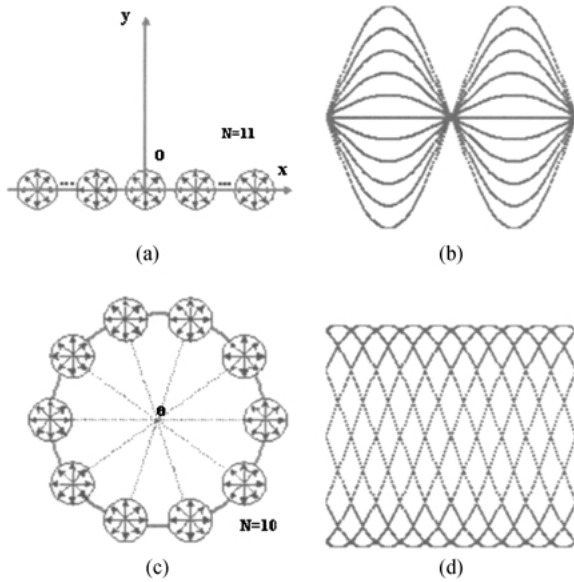


Figure 16. Hopping between panoramas: (a) along a line of 11 panoramas; (b) ray distribution of (a) in signed Hough space; (c) along a circle of 10 panoramas; and (d) ray distribution of (c).

approximating a cosine segment using a line segment. When the viewpoint is far away, the parallel mosaic approximates the perspective view very well. The re-projection or warping error increases as the viewpoint approaches the object. In addition, the image size of the parallel mosaic determines how closely the rendering camera can get to the object. Hopping using parallel mosaics and hopping using perspective images have similar warping errors, especially if constant depth is assumed.

However, rebinned parallel mosaics can have a much higher resolution than the original image if a very dense

sequence is captured. For example, we can obtain a 1D parallel mosaic of 640×240 from 640 original images with size 320×240 . Close-up views rendered from rebinned parallel mosaics have better quality than simply zooming-in the original images.

6. Experiments

6.1. Synthetic Environments

We represent a synthetic environment with 41 concentric mosaics (with size 2400×288) on 11 concentric circles. There are 21 tangent concentric mosaics, and 21 normal concentric mosaics. Note that the center mosaic degenerates to a single perspective panorama, as shown in Fig. 17(a). At the outermost circle, the tangent concentric mosaic is shown in Fig. 17(b), while the normal concentric mosaic is shown in Fig. 17(c). By hopping between these mosaics, we render five images from the left, right, center, front and back viewpoints shown in Fig. 17(d). Parallax effects (both lateral and looming) are clearly visible from the rendered images. And hopping between these mosaics provides a smooth navigation experience. However, one can only switch lateral motion and looming motion at the center. In conventional rendering with concentric mosaics, we would have used 720 such mosaics. Therefore, manifold hopping requires much less data for a similar viewing experience.

A much larger environment can be constructed by combining more mosaics captured at different locations. By carefully adjusting constant depths used for different sets, we can hop smoothly from one circle to another, and inside a circle.

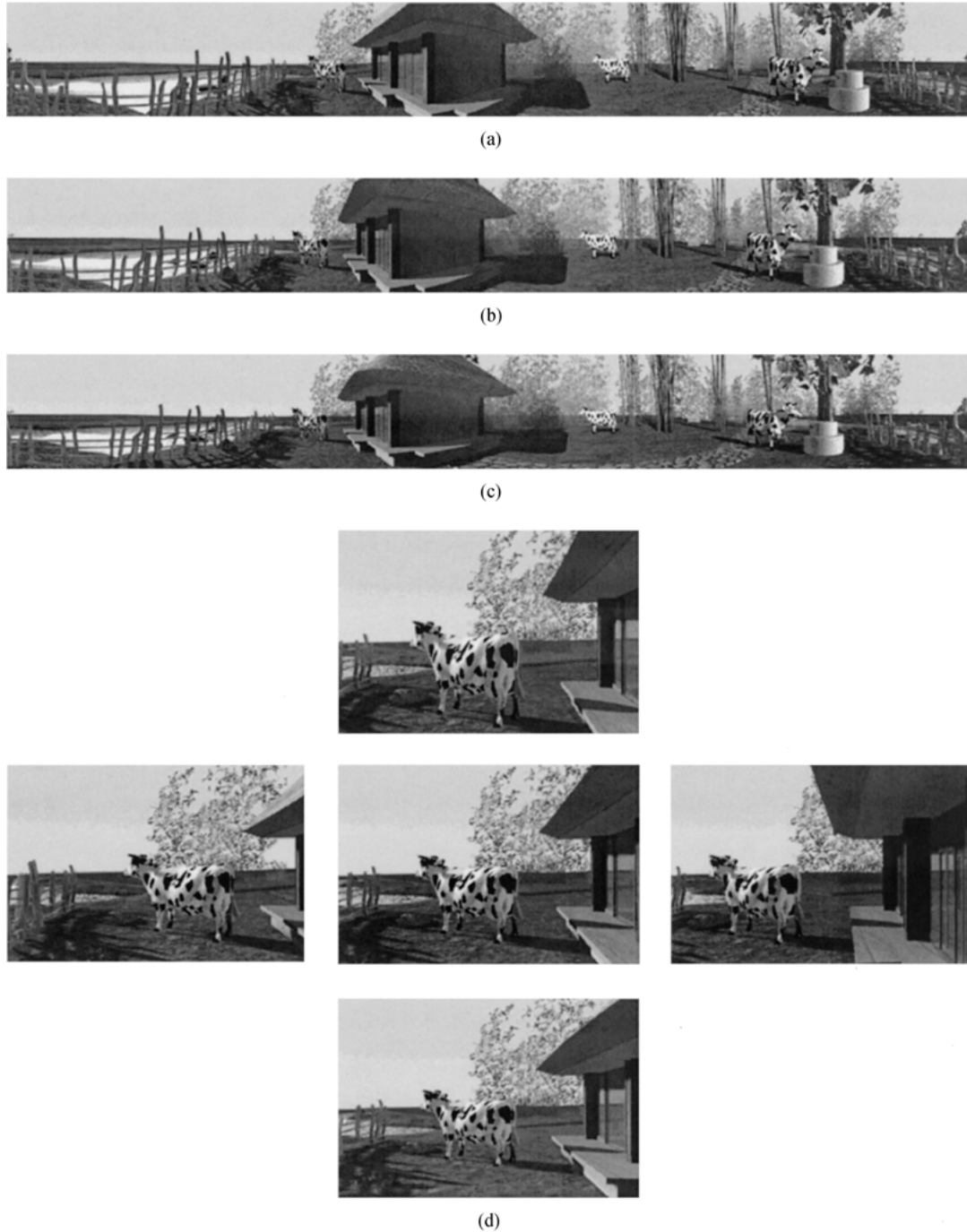


Figure 17. Lateral and looming hopping between concentric mosaics of a synthetic environment: (a) a tangent concentric mosaic; (b) the middle panorama; (c) a normal concentric mosaic; each mosaic has the size of 2400×288 ; and (d) five rendered views from manifold hopping at the left, center, right, forward and backward locations. Note that the horizontal parallax is clearly visible between the left and right views; the looming effect can be seen from the forward and backward views.

6.2. Real Environments

We have used a Sony Mini DV digital video camera to capture concentric mosaics of a real environment. The camera rotates along a circle. The video is digitized at the resolution of 720×576 . A total of 5726 frames are captured for a full circle. The raw data for the video sequence amounts to a total of 7 Gigabytes. Instead of using 720 rebinned concentric mosaics of size 5726×576 , we select only a small subset (typically 21) of resampled concentric mosaics.

Three rebinned concentric mosaics are shown in Fig. 18(a). Two high resolution images (with display size 500×400) rendered from 21 concentric mosaics are shown in Fig. 18(b) and (c). Horizontal parallax around the tree and lighting change reflected from the window can be clearly observed. Constant depth correction is used in all our experiments.

To reduce the amount of data used in manifold hopping, we can resize the original concentric mosaics. As shown in Fig. 18(d) and (e), two images with low resolution 180×144 are rendered from 11 resized smaller concentric mosaics. It is important to note that simply resizing the original 11 concentric mosaics does not generate the expected concentric mosaics. Instead, mosaics of such small size should be resampled from the original dense sequence.

We have also developed a predictive coding compression algorithm for compressing concentric mosaics with fast selective decoding and random access. As a result, the above 11 concentric mosaics can be compressed to 88k with a compression ratio of 78. Two corresponding rendered images using the compressed data are shown in Fig. 18(f) and (g).

6.3. Hopping Around Objects

We have also captured a sequence of images for an object that rotates in front of a camera. From the input sequence of 5277 images of size 360×288 , we rebin 90 parallel mosaics with size 645×288 . These parallel mosaics have an angular hopping interval of 4 degrees. A perspective image from the input sequence is shown in Fig. 19(a). Using the Lagrange interpolation, the rebinned 1D parallel mosaics are rather smooth, as shown in Fig. 19(b).

Figure 19(c) and (d) show two warped images from the 1D parallel mosaics. The parallax and lighting change can be seen very clearly in the accompanying videotape. In our experiments, we have found that

hopping angularly at an interval of 4 degrees provides a very smooth perceived camera movement. Two close-up views along the same viewing direction of Fig. 19(c) are also shown in Fig. 19(e) and (f). Because parallel mosaics have a higher resolution than the original images, close-up views provide details that would not be possible by simply zooming-in on the original images.

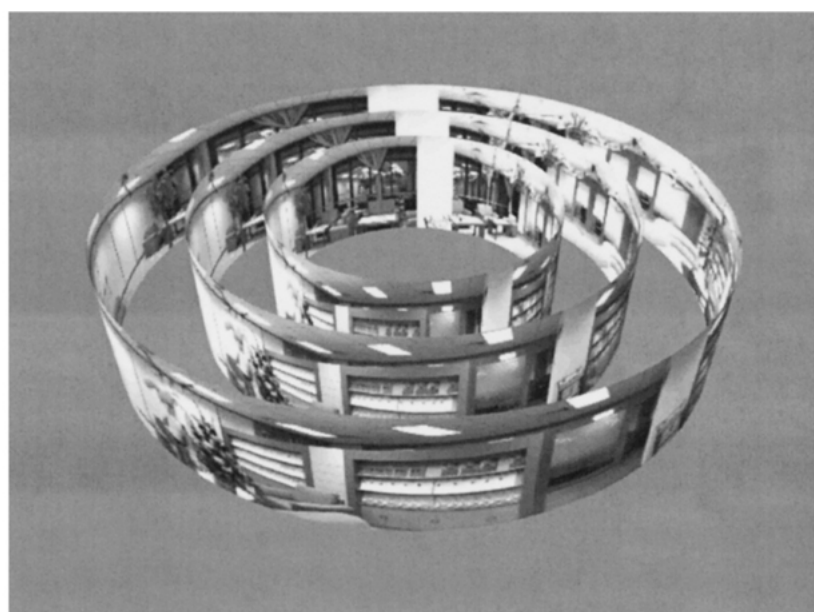
With angular hopping interval of 4 degrees in both longitudinal and latitudinal directions, we have also rendered synthetically a sequence of 2D parallel mosaics. Hopping between this collection of parallel mosaics again provides perceived smooth camera movements in two dimensions.

7. Discussion

While reducing data significantly, manifold hopping limits the freedom of user movement. In hopping with concentric mosaics, for instance, a user can only rotate along one of the concentric circles. The user is not allowed to rotate at any given viewpoint except in the center. As shown in the synthetic experiments, the user can only change from lateral hopping to looming hopping at the center. If the number of concentric mosaics is sufficiently large, it is also possible to hop around any fixed point in the angular direction by warping different concentric mosaics. In the signed Hough space, it is equivalent to finding segments from different r lines that approximate a sinusoidal curve.

Manifold hopping is not restricted to hopping with concentric mosaics or with lateral or looming movements. There are many other choices for manifolds and hopping directions. For example, hopping between panoramas has been used in QuickTime VR (Chen, 1995) using "hotspots". When panoramas are closely spaced, hopping between them can also achieve a smooth transition. Figure 16 shows two examples of hopping between panoramas. Figure 16(b) shows the signed Hough representation of a line of panoramas as in Fig. 16(a), and Fig. 16(d) shows the signed Hough representation of a circle of panoramas as in Fig. 16(c).

There are two major differences between manifold hopping with concentric mosaics and hopping with panoramas. The first difference is in capturing. Panoramas can capture similar rays to concentric mosaics as the number of panoramas increases. However, the same result will require capturing panoramas many times at different locations, as opposed to rotating the camera only once for capturing concentric mosaics.



(a)



(b)



(c)



(d)



(e)



(f)



(g)

Figure 18. Hopping between concentric mosaics: (a) three concentric mosaics projected onto cylinders; (b) and (c) two rendered images at a high resolution 500×400 ; (d) and (e) rendered images with a low resolution 180×144 ; and (f) and (g) low resolution rendered images using 88k compressed data.

The second and perhaps more important difference is in sampling. Each manifold mosaic is multiperspective, while each panorama has only a single center of projection. Since different viewpoints can be selected

as the desired path for the user, a multiperspective panorama could be more representative of a large environment than a single perspective panorama. If the multiperspective image is formed by rays taken along

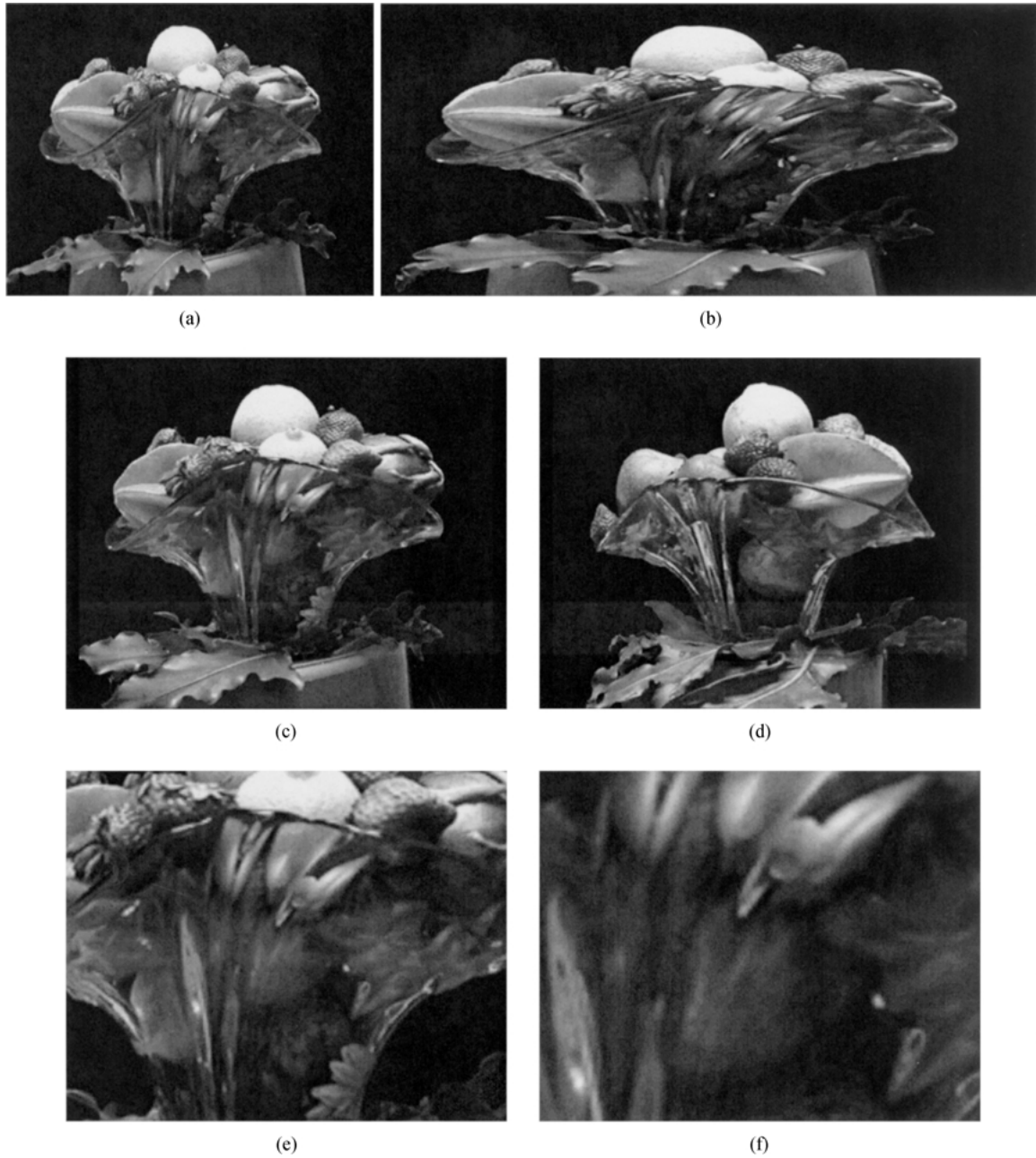


Figure 19. Hopping between parallel mosaics: (a) a perspective image from the original sequence; (b) a rebinned 1D parallel mosaic with higher resolution; (c) and (d) two rendered images from different viewing directions; and (e) and (f) close-up views along the viewing direction of (c).

the desired path of the user, the warping error from a multiperspective image is, on average, smaller than that from a perspective image (e.g., a panorama).

Concentric mosaics are suitable for the inside looking out. To observe objects from the outside looking in,

parallel mosaics can be used for manifold hopping. For concentric mosaics, the manifold is a cylindrical surface. For parallel mosaics, the manifold is a plane originating from the object center. In this paper, we have discussed manifold hopping in two dimensional space by

constraining the rendering camera on a plane. The concept of manifold hopping can be generalized to higher dimensions. The analysis in higher dimensions is very similar to the two-dimensional cases. However, it is difficult to capture such manifold mosaics in practice.

8. Conclusion and Future Work

We have described a new image-based rendering technique which we call *manifold hopping*. In summary, our technique has the following properties:

- It does not require a large amount of image data, and yet the user can perceive continuous camera movement.
- It requires neither accurate depth nor correspondence, yet generates perceptually acceptable rendered images.

Specifically, manifold hopping renders a novel view by locally warping a single manifold mosaic, without the need for interpolating from several images. We have shown that warping a single multiperspective image to a perspective image with a regular field of view causes insignificant distortion to human beings, even with warping errors resulting from incorrect depth information. Furthermore, local warping does not introduce structural errors such as double images which are perceptually disturbing.

Most importantly, manifold hopping requires relatively little input data. Capturing manifold mosaics such as concentric mosaics is also easy. By sparsely sampling the concentric mosaics, we can reduce the amount of data from the original concentric mosaics by more than 10-fold. While manifold hopping provides only discrete camera motion in some directions, it provides reasonably smooth navigation by allowing the user to move in a circular region and to observe significant horizontal parallax (both lateral and looming) and lighting changes. The ease of capture and the very little data requirement make manifold hopping very attractive and useful for many virtual reality applications, in particular those on the Internet.

Table 1 compares how manifold hopping differs from previous IBR systems, in terms of their geometric requirements, number of images, rendering viewpoints and perceived camera movement. Manifold hopping stands out in that it ensures a perceived continuous camera movement even though rendering viewpoints are discrete. It builds on the observation that a fairly large amount of viewpoint change is allowed, while maintaining perceptually continuous camera movement to humans. This observation of “just-enough hopping” for reducing image samples is, in spirit, similar to the “just-necessary effort” adopted by perceptually based techniques (Ramasubramanian et al., 1999) on realistic image synthesis to reduce computational cost. While we have experimentally demonstrated the feasibility of our

Table 1. A table of comparison for different IBR techniques: Geometry requirements, number of images, rendering viewpoints and perceived camera movement.

	Geometry	Images	Rendering viewpoints	Perceived motion
Light fields (Levoy and Hanrahan, 1996; Gortler et al., 1996; Shum and He, 1999)	No/approximate	Very large (100 ~ 10000+)	Continuous	Continuous
3D Warping (Mark et al., 1997; McMillan, 1999; Shade et al., 1998; Chang et al., 1999)	Accurate	Small (1 ~ 10+)	Continuous	Continuous
View interpolation (Chen and Williams, 1993; McMillan and Bishop, 1995; Seitz and Dyer, 1996; Avidan and Shashua, 1997)	Accurate	Small (2 ~ 10+)	Continuous	Continuous
Hopping (Lippman, 1980; Chen, 1995)	No	Moderate (10 ~ 100+)	Discrete	Discrete
Manifold hopping	No/approximate	Moderate (10 ~ 100+)	Discrete	Continuous

The citations are for reference only, not meant to be complete.

choices (e.g., 21 concentric mosaics used in most of our experiments), we plan to conduct a more comprehensive study on the psychophysics of visualization for our technique.

Note

1. By this definition, MCOP images are *not* manifold mosaics.

References

- Adelson, E.H. and Bergen, J. 1991. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*. MIT Press: Cambridge, MA, pp. 3–20.
- Avidan, S. and Shashua, A. 1997. Novel view synthesis in tensor space. In *Proc. Computer Vision and Pattern Recognition*, pp. 1034–1040.
- Chai, J.-X., Kang, S.B., and Shum, H.-Y. 2000a. Rendering with non-uniform approximate concentric mosaics. In *Proc. ECCV2000 Workshop SMILE2*, Dublin, Ireland.
- Chai, J.-X., Tong, X., Chan, S.-C., and Shum, H.-Y. 2000b. Plenoptic sampling. In *Proc. SIGGRAPH 2000*.
- Chang, C., Bishop, G., and Lastra, A. 1999. Ldi tree: A hierarchical representation for image-based rendering. *SIGGRAPH'99*, pp. 291–298.
- Chen, S.E. 1995. QuickTime VR—An image-based approach to virtual environment navigation. *Computer Graphics (SIGGRAPH'95)*, pp. 29–38.
- Chen, S. and Williams, L. 1993. View interpolation for image synthesis. *Computer Graphics (SIGGRAPH'93)*, pp. 279–288.
- Gortler, S.J., Grzeszczuk, R., Szeliski, R., and Cohen, M.F. 1996. The lumigraph. In *Computer Graphics Proceedings, Annual Conference Series, Proc. SIGGRAPH'96* (New Orleans), pp. 43–54. ACM SIGGRAPH.
- Gupta, R. and Hartley, R. 1997. Linear pushbroom cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(9):963–975.
- Kang, S.B. 1999. A survey of image-based rendering techniques. In *VideoMetrics, SPIE*, vol. 3641, pp. 2–16.
- Levoy, M. and Hanrahan, P. 1996. Light field rendering. In *Computer Graphics Proceedings, Annual Conference Series, Proc. SIGGRAPH'96* (New Orleans), pp. 31–42. ACM SIGGRAPH.
- Lippman, A. 1980. Movie maps: An application of the optical videodisc to computer graphics. *Computer Graphics (SIGGRAPH'80)*, 14(3):32–43.
- Mark, W., McMillan, L., and Bishop, G. 1997. Post-rendering 3d warping. In *Proc. Symposium on 13D Graphics*, pp. 7–16.
- McMillan, L. 1999. An image-based approach to three-dimensional computer graphics. Technical Report, Ph.D. Dissertation, UNC Computer Science TR97-013.
- McMillan, L. and Bishop, G. 1995. Plenoptic modeling: An image-based rendering system. *Computer Graphics (SIGGRAPH'95)*, pp. 39–46.
- Oliveira, M. and Bishop, G. 1999. Relief textures. Technical Report, UNC Computer Science TR99-015.
- Peleg, S. and Herman, J. 1997. Panoramic mosaics by manifold projection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97)*, San Juan, Puerto Rico, pp. 338–343.
- Rademacher, P. and B.G. 1998. Multiple-center-of-projection images. In *Computer Graphics Proceedings, Annual Conference Series, Proc. SIGGRAPH'98* (Orlando), pp. 199–206. ACM SIGGRAPH.
- Ramasubramanian, M., Pattanaik, S., and Greenberg, D. 1999. A perceptually based physical error metric for realistic image synthesis. *Computer Graphics (SIGGRAPH'99)*, pp. 73–82.
- Seitz, S.M. and Dyer, C.M. 1996. View morphing. In *Computer Graphics Proceedings, Annual Conference Series, Proc. SIGGRAPH'96* (New Orleans), pp. 21–30. ACM SIGGRAPH.
- Shade, J., Gortler, S., He, L.-W., and Szeliski, R. 1998. Layered depth images. In *Computer Graphics (SIGGRAPH'98) Proceedings*, Orlando, pp. 231–242. ACM SIGGRAPH.
- Shum, H.-Y. and He, L.-W. 1999. Rendering with concentric mosaics. In *Proc. SIGGRAPH 99*, pp. 299–306.
- Shum, H.-Y. and Szeliski, R. 1999. Stereo reconstruction from multiperspective panoramas. In *Proc. Int. Conf. Computer Vision*, pp. 14–21.
- Wood, D.N. et al. 1997. Multiperspective panoramas for cel animation. In *Computer Graphics Proceedings, Annual Conference Series, Proc. SIGGRAPH'97* (Los Angeles), pp. 243–250. ACM SIGGRAPH.
- Zheng, J.Y. and Tsuji, S. 1990. Panoramic representation of scenes for route understanding. In *Proc. of the 10th Int. Conf. Pattern Recognition*, pp. 161–167.
- Zorin, D. and Barr, A. 1995. Correction of geometric perceptual distortions in pictures. *Computer Graphics (SIGGRAPH'95)*, pp. 257–264.